Scientific Theory and Possibility

Sam Baron,* Baptiste Le Bihan,† & James Read‡

Abstract

It is plausible that the models of scientific theories correspond to possibilities. But how do we know which models of which scientific theories so correspond? This paper provides a novel proposal for guiding belief about possibilities via scientific theories. The proposal draws on the notion of an effective theory: a theory that applies very well to a particular, restricted domain. We argue that it is the models of effective theories that we should believe correspond, at least in part, to possibilities. It is thus effective theories that should guide modal reasoning in science.

0. Contents

*s.baron@unimelb.edu.au
†baptiste.lebihan@unige.ch
‡james.read@philosophy.ox.ac.uk

1. Introduction

The question of what is possible and what is not is at the heart of all philosophy and has been approached in various ways from different directions. The aim of this article is to contribute to the empiricist tradition in the philosophy of modality, which attempts to discover genuine possibilities in our scientific theories—and not in a faculty of *a priori* conceivability (see, e.g., Berenstain and Ladyman 2012; Ruyant 2019; Wilson 2020). Following this approach, and looking at a scientific theory, it is very natural to suppose that its models, here regarded as solutions to its equations, correspond to possibilities with, if the theory is to be empirically adequate, at least one model corresponding to what is actual. Van Fraassen articulates this idea clearly, when he writes that:

> To believe a theory is to believe that one of its models correctly represents the world. You can think of the models as representing the possible worlds allowed by the theory; one of these possible worlds is meant to be the real one. To believe the theory is to believe that exactly one of its models correctly represents the world (not just to some extent, but in all respects). (van Fraassen, 1980, p. 47)

Physical possibility, or scientific possibility (for present purposes, we take the two expressions to be synonymous) is, we assume, a mind-independent, objective modality.[1] Van Fraassen's approach gives rise to an important question: if physical possibilities are to be found in theories' models, exactly *which models* of *which theories* should we believe correspond to genuine possibilities? For one sense of 'possibility' the answer appears straightforward. If possibility means logical possibility, then presumably a model need only be consistent for it to correspond to a genuine possibility. We should thus believe that all consistent models of all theories correspond to logical possibilities, where 'correspond' here is understood as in the quote above: as the representation of worlds by models.

For a more restricted notion of physical possibility, however, the answer is less obvious. Metaphysically speaking, the physical possibilities are those situations that are consistent with the actual laws of nature (Earman 1986, p. 13; Maudlin

[1] One might interpret the expression 'physical possibility' in a different sense, as referring to possibilities specific to physics (as opposed to those found in other sciences). However, that is not the intended usage of the term in this work. Here the physical possibilities are the ones associated to the physical (i.e., natural) world.

2007, p. 18).[2] It is thus tempting to think that all theories that accurately capture the laws have models that correspond to physical possibilities (exactly what it means for a theory to accurately capture the laws is an issue to which we return below). As we will argue here, however, this way of specifying physical possibilities via the actual laws of nature faces serious problems. We thus suggest an alternative approach based on the notion of an *effective theory*, which aims to capture *effective laws*: general principles that apply only to a specific domain. We use those terms as generalisations of the way they are used in the context of effective field theories where 'specific domains' are identified to energy scales.[3] It is effective theories whose models we should believe correspond to physical possibilities.[4]

The paper is structured as follows. We begin, in §2, by identifying four desiderata that shape the search for a belief-guiding principle of the right type. We then show why simple law-based approaches won't do, since they don't satisfy the desiderata. In §3 we outline our own principle based on effective theories. In §4 we address some potential difficulties with our approach. In §5 we wrap up.

2. Four Desiderata

Our goal is to identify a principle for guiding belief. In particular, a principle that tells us which models of which theories we should believe correspond to genuine possibilities. Why do we need such a principle? The short answer is that modal reasoning is deeply important to science. Consider physics. As Wilson puts it:

[2] For present purposes, we will simply assume realism about the laws of nature. However, while laws play a role in our account, we won't assume any particular theory of laws. We don't think realism is essential to our position, but this is not an issue we will pursue here. We also assume a kind of minimal realism about physical modality compatible with a variety of radically different approaches (such as, e.g., Cartwright 1983 and Berenstain and Ladyman 2012).

[3] Although we borrow the term from the context of effective field theories, our purpose is tangential to the recent discussions of the relation between effective field theories and scientific realism (Williams 2019; Rivat 2021). We use the label 'effective' to refer to the fact that all empirically confirmed scientific theories are effective in a broader sense—i.e., one which does not reduce the notion of validity domain of theories to the notion of energy scale.

[4] To contextualise our work: we regard our project here are being complementary to e.g. the discussion of 'situated possibilities' by Ruyant (2021). The central innovation which our own work brings to the table is the notions of an effective theory and of an effective law—notions which have not so far been put to work in the literature on science and modality.

> It is a truism that physics involves reasoning about alternative physical possibilities. To explain and predict actual observations, we construct models of the phenomena which represent alternative physically possible histories; perhaps we also assign a probability distribution over these histories. This essentially modal character of physics has been emphasized by authors as varied as Sellars, Suppes, Cartwright, Ladyman and Ross, and Maudlin. (Wilson, 2021, p. 1113)

Here are a couple of examples. First, symmetry explanations. Some explanations appeal to the fact that certain features of a physical system are invariant under some salient transformation. For some, the fundamental quantities of physics just are those that remain invariant under changes, whereas non-fundamental quantities are variant (see e.g. Saunders 2003).[5] This notion of fundamentality forges a tight connection between physical and modal notions since an invariant quantity is one that is necessary in the sense that it appears in all the possible worlds related by the transformation. This makes it quite important to ensure that we are working with the *right* possibilities when providing such explanations.

Second example: statistical mechanics. According to the popular 'Boltzmannian' approach, one can characterise entropy as follows: the (Boltzmann) entropy of a macrostate is a measure of the number of possible microstates with which it is compatible. This Boltzmannian notion of entropy clearly makes use of a notion of possibility, one that it seems natural to call physical possibility. Indeed, the alternative 'Gibbsian' approach to statistical mechanics countenances ensembles of physical systems that evolve in a probabilistic way, and thus would also appear to trade in notions of physical possibility. Again, we need to ensure that we attend to the right physical possibilities when reasoning in this area, and for that some epistemic guide to possibility is needed.

These examples draw from physics, but of course examples can also be found elsewhere. Take, for instance, Sober (1983)'s equilibrium explanations, which appeal to the instability of certain possibilities to explain actual facts. Sober's classic example is the explanation of the 1:1 sex ratio in most animals, which relies on the fact that all other possible ratios are unstable and tend toward 1:1 given enough time. Indeed, on what is perhaps the most prominent account of

[5] Think also of Noether's theorems: conserved quantities are associated with symmetries of an action. See Read and Teh (2022) for recent work on the philosophy and physics of Noether's theorems.

explanation, all scientific explanations outside of physics rely on counterfactuals (see e.g. Woodward 2003). That being so, we should assume that possibilities play a large role in science across the board and so, again, there is pressure to identify a reliable method for guiding belief about physical possibility so that we can ensure the explanations provided are good explanations.

As we see it, a viable principle for guiding belief about possibility should tell us which situations we should believe are genuine possibilities, and should moreover satisfy the following four desiderata:

- D1 Provide a basis for reliable modal reasoning.
- D2 Provide a guide to possibility now, rather than at some putative 'end' of science.
- D3 Allow some superseded theories to guide beliefs about possibility.
- D4 Be generalisable across science.

For a principle to satisfy D1, any situations the principle tells us to believe are genuine possibilities should in fact be genuine possibilities. That is, the principle should not just guide belief but it should do so accurately. For a principle to satisfy D2, it must be applicable now, to our current scientific theories. It should not, for instance, recommend agnosticism about what the possibilities are in lieu of a final theory of everything.

Next, D3. By a superseded theory we mean any theory that scientists consider to have been replaced by a better theory. One example is Newtonian gravitation. This theory has been superseded by general relativity. It is thus tempting to think that the models of Newtonian gravitation don't correspond to genuine possibilities. That, however, looks like a mistake. Consider, for instance, the solar system: this system is modelled well by Newtonian gravity, and indeed different models of Newtonian gravity can provide possible ways for regions of the world like this (i.e., other stellar systems, of which there are of course many!) to be. Given this, we want an epistemic principle that is flexible enough to allow for Newtonian gravitation to guide belief about possibility in some, but by no means all, situations.

Finally, D4. When we talk of physical possibility we don't simply have in mind *physics* possibilities, though that is a large portion of the types of possibilities at issue. Rather, by physical possibility we mean to include possibilities for the physical world in general, so long as these are ratified by science. We think it would be odd indeed to have one method for guiding belief about possibility in one area of science, and a different method for another area of science. This would make our epistemic access to possibilities through science unattractively dis-unified. That being said, we grant that this might be forced on us, if different areas of science somehow call for different epistemic access to modality. But we think it is reasonable to assume unification as the default position, and try to fit a principle on this basis.

With these four desiderata in hand, we turn now to a simple law-based principle for guiding belief about possibility. Here's the principle:

> For any theory that accurately captures the laws, every model of that theory corresponds to a physically possible world.

On the face of it, this principle satisfies our first desideratum beautifully. For what it invites us to do is believe in the possibilities rendered by theories that accurately capture the laws. Any such theory, one might think, is bound to have models that correspond to genuine possibilities, given that those possibilities just are circumscribed by the very laws captured by the theory.

Unfortunately, the principle fails to satisfy D2 because it is doubtful that *any* of our past or present scientific theories manage to accurately capture the actual laws of nature. First: what is it for a theory to capture the actual laws? A law of nature, presumably, is an exceptionless, non-accidental regularity that applies universally. A theory can be regarded as a class of models.[6] This class of models can generally be described using a set of equations or general principles. A theory accurately captures the actual laws of nature when the sets of equations or general principles correspond to the actual laws. In other words, the actual laws, when written down, just are the equations or general principles of a given theory that defines a class of models.

Now, while many of our theories imply universal claims that are candidates to be laws, none of them imply such claims that are true. The generalisations that

[6] Here, we sidestep the debate between semantic versus syntactic formulations of physical theories. In light of the work of Lutz (2017), it is not obvious that this is a substantive debate in any case.

we find in even our best theories neither hold universally nor lack exceptions.[7] Indeed, it seems plausible that it is only at the end of science that our theories will accurately capture the laws. That being so, the law-based principle stated can only guide belief when science is done.

This, one might argue, is easily fixed. For while our current theories are not fully accurate, insofar as they don't capture the actual laws, they approximate the actual laws to a high degree.[8] We can capture this thought by weakening the above principle as follows:

> For any theory that approximately captures the laws to a high degree, every model of that theory corresponds to a physically possible world.

The approximation principle arguably satisfies D2 (assuming we can establish that our current theories approximate the laws to a high degree). It also seems to satisfy D1. On the assumption that our theories closely approximate the truth, our beliefs about possibilities based on those theories will be good approximations to the truth as well. However, the principle seems to fail D3. The reason for this is that there appear to be superseded theories that do not even approximately capture the laws, and yet are good guides to possibility.

Take, for instance, Newtonian gravitation. This theory works under specific circumstances which don't correspond to the majority of gravitational situations that are, for instance, covered by relativity. Newtonian mechanics is also not a quantum mechanical theory. This means that a vast number of quantum gravitational situations lie outside the domain of the theory. For a theory to approximately capture the laws, in this case the laws of gravitation, it should apply well to

---

7 We are not the first to point out that our scientific theories fall short of accurately capturing the actual laws of nature, whatever they may be. Both Cartwright (1999) and Lange (2012) offer similar examples. In both cases, they point to apparently universal claims, and show that they have exceptions. Cartwright, for instance, focuses on Newton's law of gravitation, and shows that it applies only under certain *ceteris paribus* conditions (e.g., at relatively low speeds, or in the absence of electromagnetic interference). Lange, by contrast, considers general principles of thermal expansion found in thermodynamics, and shows that these principles apply at least conditionally. For instance, general principles of thermal expansion apply to metal rods only in the absence of external factors influencing the system, such as mechanical impact on the rod (e.g., someone hammering the rod into place).

8 Note that the approximation here doesn't involve any domain restriction. The idea is that our theories capture the entirety of physical possibilities, approximately. Our proposal later on uses approximation, but under domain restriction: we focus on laws that hold approximately for a restricted domain, rather than an entire physical possibility.

most gravitational situations. Accordingly if what we are looking for is a theory that approximately captures the actual universal laws to a high degree, then Newtonian gravitation is not a viable candidate. Nonetheless, and as already noted, Newtonian gravitation is still used for a variety of purposes, precisely because it seems to give us good information about physical possibilities.

Is there a version of the law-based approach that does better? We doubt it. The trouble is that such principles generally set the bar too high. For such principles, belief about possibility is guided by the most general laws and principles. But this will always pull against cases in which superseded theories seem to accurately guide modal beliefs, since those theories are generally superseded on the grounds that they fail to be fully general, or to even be close. Such principles, simply put, start in the wrong place. Rather than starting from the most general laws, we should focus more on the success of theories in specific situations, using that as a guide to belief about possibility instead and then generalise if we can. It is to a principle along these lines that we now turn.

3. Effective Theories

Our suggestion is to focus on *effective laws*. An effective law is not a universal, exceptionless regularity. Rather, an effective law is a regularity that (i) holds for a specific domain D and (ii) breaks down for a specified domain $D^* \neq D$.[9] Domains are specified in terms of the conditions that hold within them. For instance, a region in which the gravitational field is very weak, and everything is moving at low speeds relative to the speed of light, corresponds to a particular domain. Similarly, a particular energy scale corresponds to a domain, such as the extremely high energies at which general relativity is thought to break down, or the comparatively low energies at which general relativity applies.

Effective laws are grounded in universal laws. Which is to say that the universal laws plus the conditions which specify a given domain ground the truth of the effective law for that domain. An effective theory is a theory that captures the effective laws for a domain either exactly or to a high degree of approximation. Our proposal, then, is to focus on effective theories.

[9] Note that from this it follows that no final theory can supply effective laws, assuming that final theories are unconstrained in application domain-wise. If one wishes to include final theories as the limit case of effective theories, one could lift the second condition, but that would make it harder to characterise an effective law.

Because effective theories apply only to a specific domain, we cannot have a totalising principle like that considered in the previous section. That is, a principle that provides a link between models and entire worlds. Instead, we need a way to link the models to possibilities in a more specific sense. Here we draw on similar recommendations made by a number of philosophers to focus on parts of worlds, instead of entire worlds (Kripke 1981; Ladyman and Ross 2007; Lange 2016; Ruyant 2019).

Effective laws thus correspond to only a *part* of the world, and thus not to the world as a whole and so the models of effective theories approximate parts of genuine physical possibilities. With this in mind, we offer the following principle:

> For any model M of an effective theory T, M approximates a part of a physically possible world.

We have strong evidence that some of our current theories capture actual effective laws. Take, for instance, general relativity. The principles of gravitation from general relativity don't hold universally, for all domains. However, we have good evidence that they hold universally for a particular domain, namely certain energy scales. The same point applies for the Standard Model of particle physics, for Newtonian gravitation, or indeed for essentially any other extant empirically successful theory of physics.

Given that we have strong empirical support for effective theories, it is reasonable to believe that they capture effective laws either exactly or to a high degree of approximation. Therefore, it seems we can rely on such theories in our modal reasoning, at least insofar as we are reasoning modally about the domain of applicability for each theory, and thus what we are doing is drawing inferences about parts of worlds (as per our principle). For that reason, our principle satisfies D1. D2 is also satisfied: the whole point of effective theories is that we have them now, not at the end of science. Thus we can use them to draw reliable modal inferences without having to wait until the discovery of some putative final theory.

The great power of our principle, however, lies with its capacity to satisfy D3. Take the case of Newtonian gravitation again. Even though this theory might not provide a very good picture of the actual universal laws, as already mentioned it does provide an accurate picture of effective laws for a specific domain. In particular, for domains in which the gravitational field is weak and velocities are low, Newtonian gravitation is highly accurate, and very well-supported empirically. Moreover, we can be confident that new data won't overturn this fact, and so we can be confident that Newtonian gravitation is accurately capturing effective laws. It is for this reason that its models correspond to physical possibilities, albeit approximately and only in the sense of corresponding to parts of possible worlds.

Now, we recognise that there is perhaps a bit more to say here. We are claiming that the empirical resilience of an effective theory is a sufficient basis upon which to conclude that the effective laws will continue to hold within a domain. One might object, however, that empirical resilience is not enough. We also need some kind of robustness condition that provides a safeguard against defeaters for effective laws in a domain. We will return to this issue later on, in §4.1.

This brings us to D4. To satisfy D4, our principle must be generalisable across science. On the face of it, it generalises beautifully. Take biology, for instance. It is plausible that there are effective laws of biology: laws that apply to certain domains (such as for all biological organisms) that are ultimately grounded in the universal laws at some level. As such, our principle applies equally well to biological theories, or indeed to any effective theories in any domain, as to theories from physics. For extremely well-confirmed theories of any domain where we can be confident that new data won't undermine the theory, we can be relatively certain that the models of the theory approximate parts of physical possibilities. For theories that are less well-confirmed, we can be less confident of the modal inferences drawn from those theories. But that is as it should be: it seems right for theories that enjoy less evidential support to be less of a guide to what is possible.

Because effective theories can be highly specific, we find it hard to see that many theories will be left out of the picture, regardless of which particular area of science they might come from. For keep in mind that effective theories must be very well-confirmed for their domain. Indeed, the level of confirmation is quite demanding: we must be fairly confident that no new data will require significant revisions to the theory. Any theory like that, no matter how specific, will count as an effective theory.

Thus, the only theories that our principle won't capture will be theories that are not well-confirmed empirically. But we don't see that as a problem. For, again, we really should be careful about drawing inferences from such theories (modal or otherwise), and so it seems reasonable to be sceptical that the models of these theories correspond to physical possibilities.

We are claiming that strong accord with empirical data is good evidence that a theory is an effective theory: one that approximates, to a high degree, the effective laws for a particular domain. Given this, why not just say that strong accord with empirical data is good evidence that a theory accurately captures the actual universal laws, not just the actual effective laws? The answer is that in the case of effective laws, we can be confident that there is very little, if any, extra empirical data to be gathered that would undermine the accuracy of the theory for a given domain.[10] Thus, we have good reason to suppose that the principles articulated by the theory won't need to be modified much if at all to capture the actual effective laws. In the case of a theory and the universal laws, however, we have no similar assurance. There may well be a great deal of empirical data that would require us to change the general principles articulated by a theory in order to get it to match the actual laws.

Take general relativity again. This theory does not capture fully the actual laws, as already explained. We know that there are facts about high-energy domains that will need to be accommodated. We currently have no idea about what those facts might be or about the scope of the revisions to general relativity which might be required in order to bring it into line with the actual laws. As a result, we have significant uncertainty about how close general relativity is to capturing laws of gravitation. If, however, we focus just on the low energy domain, we know that most of the data relevant to general relativity in this domain have been gathered already. Thus while general relativity might not get the effective laws for this domain exactly right, we can be comparatively confident that it is quite close. In particular, we can be confident that no new information about this domain will come in and radically undermine the theory.[11]

The shift from correspondence with possible worlds to correspondence with parts of worlds might seem troubling. Granted, we are able to reduce uncertainty about whether the parts of worlds at issue are parts of physical possibilities, but we do so apparently at the cost of our capacity to draw reliable inferences about complete worlds. We admit as much: our principle provides no way to work out which models of which theories correspond to complete worlds. But we don't see this as a problem. For we rarely if ever use science to draw modal inferences about worlds as a whole. We are typically interested in possibilities in a narrower sense. That's because, when we draw modal inferences from science, we generally draw those inferences with respect to specific domains. We are therefore confident that our principle is sufficient to guide beliefs about which models of which theories correspond to physical possibilities, at least insofar as we want those beliefs to be both reliable and useful in modal reasoning within science.

---

[10] See Hoefer (2020): "*the experiments have all been done already, and current theory is confirmed by them*" (emphasis in original).

[11] Leading on from the previous footnote: this is a point which has been made with great force by Hoefer (2020).

4. Problems

So far we have introduced a new principle that aims to determine which models of which theories correspond to physical possibilities. In this section we consider and respond to some potential problems.

4.1. Not Every Model

The principle, as stated, is quite general: *every* model of an effective theory approximates part of a physically possible world. This, one might argue, is a problem. Even for effective theories, one might argue, we should allow that some of their models don't correspond to any genuine physical possibilities. Here are two examples from the philosophy of spacetime in particular:

i. Those inclined towards relationalism about spacetime might aver that vacuum solutions of general relativity do not represent physical possibilities, for there can be no spacetime in the absence of matter. (Pooley 2001, for example, explores this idea.)

ii. One might maintain, *à la* Earman (1995), that solutions of general relativity which admit of closed timelike curves and which violate certain consistency conditions (designed to rule out grandfather-type paradoxes) do not represent physical possibilities.

These kinds of cases can be handled, to some extent, using the notion of approximation already built into our principle. As noted, the models of an effective theory correspond to parts of physical possibilities by approximating them. This means that there can be some measure of mismatch between the models and the physical possibilities. This mismatch leaves it open that some aspects of a model don't correspond to anything in a given physical possibility.

This is, perhaps, what is happening in case (ii), if we follow Earman. Solutions to general relativity which feature closed timelike curves would only approximate physical possibilities, since they get the topology wrong. One way in which they might get the topology wrong is through the inclusion of closed timelike curves. Nevertheless, such solutions might still approximate physical possibilities to some degree, for (say) the local physics described in such solutions might still agree with the local physics in (a region of) a world devoid of closed timelike curves. This is related to e.g. the content of Proposition 1 of Manchak (2016), which states that any spacetime is locally observationally indistinguishable from a spacetime which lacks a global time function. Moreover, recall that on the account we are proposing, it is only parts of worlds that models of effective theories need to approximate, and so the case of closed timelike curves can be handled. To be clear: we are not saying that only parts of models represent parts of worlds. The representation relation is between the whole model and parts of a world, it is just that the relation need not be exact, it can be approximate, leaving room for the model to get the physics right despite including aspects that are not physical (such as closed timelike curves).

Case (i) is perhaps a little harder. Consider again relationalists about general relativity who maintain that at least some of the models of the theory don't approximate even parts of physical possibilities.[12] If this can happen, then uncertainty about the reliability of our modal inferences starts to creep back in. For if there are some models that don't approximate any physical possibilities at all, then how confident should we be that any particular model approximates a physical possibility?

In order to address this issue, we propose the following modification of our principle: we should believe that all models of an effective theory approximate physical possibilities, in the absence of countervailing physical principles for which we have strong inductive evidence based on past theories.[13] The thought, then, is that the models of an effective theory are innocent until proven guilty: our default position is to assume that they all approximate physical possibilities.[14] This default position, however, is subject to specific defeaters, though what those defeaters might be is considered on a case-by-base basis.

Given this modification we can thus allow that some models of an effective theory don't correspond to even parts of physical possibilities. The idea is that defeaters 'knock out' specific models, without impugning the entire set of models associated with an effective theory. This helps us to handle the case of relationalism, since we can allow that specific models don't correspond to parts of physical

[12] One could push back against this, by e.g. arguing that vacuum solutions still represent *parts* of physical possibilities—e.g., Minkowski spacetime can still represent parts of solutions to e.g. Maxwell's equations. In which case, the worries and responses to follow are moot. Consider also the point made by Lehmkuhl (2017), that vacuum solutions to general relativity can still represent physical (sub)systems (e.g., the vacuum Schwarzchild solution can represent the solar system) when interpreted 'carefully' rather than 'literally'.

[13] We are grateful to a referee for suggesting this modification.

[14] What we propose is in the spirit of (a modal version of) the 'totalitarian principle', popularised in physics by Gell-Mann (1956): "everything not forbidden is compulsory".

possibilities.

What if there are always some defeaters? Then we can weaken our principle slightly by including a second dimension of approximation. We can thus add that the space of models is itself an approximation to a space of models that captures the physical possibilities exactly. A weakened principle along these lines can be stated as follows:

> For most models M_i of an effective theory T, each of the M_i approximates a part of a physically possible world.

This weakened principle reduces the reliability of modal inferences drawn on the basis of science, since it introduces a second source of error. The first source of error consists in the fact that models now merely approximate parts of physical possibilities. The new source of error is based on the fact that some models might not approximate parts of physical possibilities at all. While this second source of error reduces the reliability of modal inferences drawn using science, it doesn't undermine them completely. For it is still the case that most models of effective theories approximate parts of physical possibilities, and so our inferences drawn using those theories are still generally reliable. Moreover, one can still get a handle on the subset of models which (approximately) represent (parts of) physical possibilities, once one is explicit about the above-mentioned countervailing philosophical principles in play. That is still enough to capture the importance of modal inference to science.

In this section we have proposed a modification for handling specific defeaters: reasons that provide a basis for denying that some model or other of an effective theory corresponds to a physical possibility. We can use a similar modification to assuage a related concern identified in §3. There we claimed that empirical resilience—being supported by data, and unlikely to be overturned by new data—is sufficient grounds for believing that an effective theory will continue to hold in a given domain. As discussed, one might worry that this is too weak, since there could be other defeaters that undermine the effective theory that are not ruled out by empirical resilience. What we can do is write the above modification into our notion of an effective theory. An effective theory is one that is empirically resilient and not subject to specific defeaters stemming from any currently known countervailing physical principles that have a strong inductive basis in past theory. This helps to make the notion of an effective theory more robust, since it will ensure that effective theories are those theories that aren't somehow undermined by implications from other theories (rather than by empirical violations in their domain of validity).

4.2. Domain Validity

Our picture relies on the idea that effective theories are theories that capture effective laws, which apply to restricted domains. One might worry, however, that the domain of validity of our theories is something about which we can know only in retrospect. For instance, it could be argued that we knew the exact domain of validity of Newtonian gravitation only once we developed relativity. Thus, there could be unsuspected domain violations lurking for our current theories. If that's correct, though, then we might not have a good grip on which of our theories are effective theories, and perhaps might not really know until the development of a final theory. The capacity of our account to satisfy D2 starts to look shaky.

In response, we believe it is useful to clarify what we mean by a 'domain of validity'. By a domain of validity we just mean the set of circumstances under which a theory is empirically adequate. We take it that one could know the domain of validity of Newtonian gravitation in this sense prior to the development of (say) general relativity: we can simply look at the situations under which Newtonian gravitation is empirically adequate and read the domain off those situations. There might be other ways to specify the notion of 'domain of validity' but for present purposes we set these aside.

What this means is that we do not identify effective theories by finding some domain, and then looking for a theory that is valid in that domain. Rather, an effective theory is one that works well for a set of circumstances, namely just that set of circumstances for which it is empirically adequate. In this way, an effective theory's domain of validity and its empirical adequacy are packaged together. Thus, we see no problem in satisfying D2 arising from the notion of a domain of validity. All we need to do is identify theories that are empirically adequate for a set of circumstances, resilient across those circumstances in the sense that we've tried to find violations in those very circumstances and haven't found any, and that are not subject to known defeaters stemming from other theories.

Now, one might remain worried. Empirical resilience and protection from defeaters gives an effective theory some degree of robustness, but perhaps it doesn't entirely immunize an effective theory from theory change. That is, in cases of theory change it might be that an effective theory gets overturned.

But we don't see that this is right. Theory change doesn't undermine a theory's capacity to capture a range of empirical data. What generally happens is that we move to new theories that capture a broader range of data. Again, we suspect that any worries here about theory change are related to how we think about the domain of validity. Suppose we take a domain of validity in the second manner suggested above, as a set of features we want to capture antecedent of any theory. We then fit a theory to that domain. If this is what we mean by an effective theory, then theory change poses a problem: it may be that a new theory will capture that domain better, and so be a better effective theory, in the sense of a theory that captures this pre-specified domain very well.

But, as before, that's not what we mean by a domain of validity. We just mean the circumstances under which a theory is empirically adequate, where we've looked for violations in those circumstances and haven't found any, making the theory empirically resilient in those circumstances. Theory change doesn't depose effective theories specified in this way, because it doesn't stop a theory from being empirically adequate in the circumstances within which it is empirically adequate.

Finally, it's worth noting that our approach doesn't require that we be absolutely certain our theories are effective theories. We require only that this belief is justified. That being so, we will set aside the task of trying to provide an apodictic basis for believing that an effective theory suffers no as-yet-unknown violations for its domain of validity. All we need, then, is a strong basis for believing that some theories are in fact effective theories. For three reasons, we think that such a basis can be found.

First, the massive empirical success of the theory in a given domain. This gives us confidence that the theory has captured a wide range of data. Second, the lack of any detected violations of the theory, where many attempts have been made. This gives us confidence that there aren't any hidden violations. That's simply because we have looked for such violations and have failed to find them significantly often. Third, the lack of any countervailing principles inductively supported by past theories that would indicate violations in the current domain of validity for a theory. Thus, we can acquire further confidence that a theory is an effective theory by drawing on the robustness condition offered in §4.1.

Taken together, these three considerations don't *guarantee* that there are no hidden violations in a domain of validity for a given theory—but, again, we do not require any such guarantee. We need only sufficient confidence to be able to identify effective theories, and use them as a guide to possibility.

4.3. Approximation

Our principle uses a notion of approximation. However, we haven't said much about what this is. One might object, then, that our principle is intolerably vague.

As we see it, there are at least three ways in which models might 'correspond' to worlds:

1. A model corresponds to a world if it is isomorphic to a complete possible world.[15]
2. A model corresponds to a world if it is isomorphic to a part of a complete possible world.
3. A model corresponds to a world if it represents either a part of a complete possible world or a complete possible world to some degree.

Our principle uses the third, weakest notion of correspondence. Thus, when we talk of *approximation* we mean *representation to some degree*. What is it for a model to represent to a degree? It depends on how models represent.[16] This is a difficult issue, and not one we aim to discuss here for there are many accounts available. The point is just that whatever one's account of the representation relation, it is a desideratum on such accounts that they be able to make sense of more or less accurate representations. That, however, is all we need for approximation relations between models and (parts of) worlds.

If approximation is understood in terms of degree of representation, then a further question arises: to what degree should we believe that the models of an effective theory represent and thus approximate physical possibilities? Our answer is that we should believe that the models of an effective theory approximate physical possibilities to a high degree, in the absence of any countervailing considerations (of the kind that might lead us to believe that the models of a theory don't approximate physical possibilities at all, as discussed above).

What justifies this claim? The claim is justified by the fact that the models at issue are models of an effective theory, which is highly accurate for a domain (as defended in §4.1-4.2). When a theory manages to capture the effective laws for an actual system, it seems plausible to suppose that its models also do a good job of capturing possibilities for that system (again, in the absence of reasons to think otherwise). In short, the actual is a good guide to what is possible. We admit that this connection between actuality and possibility could be rejected. If one rejects the connection, however, then it is unclear why we should take any theory to be a guide to possibility in any sense. Rejecting the connection would thus lead to a version of modal scepticism about physical possibility. Since we reject scepticism

15 There are serious questions regarding what it could mean for models to be 'isomorphic' to worlds—see e.g. van Fraassen (2002). We register these as serious and set them aside.

16 See Frigg and Nguyen (2017) for an overview of various approaches.

of this kind, we take it to be a reasonable assumption that the models of effective theories approximate physical possibilities to a high degree.

4.4. Effective and *Ceteris Paribus* Laws

Thus far we have employed the notion of effective laws. One might argue, however, that effective laws are nothing but *ceteris paribus* laws. The *ceteris paribus* conditions on laws, however, are notoriously difficult to spell out and, in the end, may collapse into triviality. There is thus a real risk that our account will inherit these problems.

Effective laws are different from *ceteris paribus* laws, however. A *ceteris paribus* law is generated by taking a universal exceptionless regularity and adding extra conditions that constrain the application of the law. An effective law, by contrast, is not a way of modifying an existing universal law with extra conditions. It is a non-universal law that holds for a specific domain. The effective laws are grounded in the universal laws, whereas *ceteris paribus* laws are supposed to be replacements for universal laws.

Unlike *ceteris paribus* laws, effective laws are not difficult to spell out. In order to spell out an effective law, we simply identify a particular domain of application, and then state a general principle for that domain. Since we can specify domains of application by identifying properties of physical systems (such as energy scales, gravitational strength, and so on) there is no problem with stating effective laws. There is thus no reason to suppose that the problems that plague *ceteris paribus* laws will apply to effective laws as well.

4.5. Grounding of Effective Laws

We have claimed that effective laws are grounded in universal principles. We have also said that effective laws can be highly specific. One might object, however, that these two points are in tension. Take the following example. Suppose we come up with an effective theory about the breeding cycles of finches, which accurately captures an effective law. It's implausible, one might argue, to suppose that this effective law is grounded in a set of universal laws. The universal laws, whatever they might be, will be far too general, and the effective finch law will be far too specific. Effective and universal laws would involve too different sorts of things, jeopardising the possibility of grounding the former in the latter.

We admit that there might be some implausibility to the idea that the finch law is *directly* grounded in the universal law *alone*. But, firstly, effective laws can, and likely will be, indirectly grounded in universal laws, via legal intermediaries. In particular, each effective law will be grounded in another effective law that is perhaps more general, applying to a domain that is broader than but includes the domain covered by the highly specific effective law. Indeed, as we see it there will be a chain of effective laws that issue from the most basic, universal laws, whatever they might be to the most specific effective laws. The models of theories that capture effective laws at each step of the chain will correspond to physical possibilities, as and when they are developed.

Secondly, that effective laws are grounded in universal laws doesn't mean that universal laws are their *sole* ground. Rather, effective laws are *partially* grounded in the universal laws. Other factors contribute to the grounding of effective laws as well: the concepts used to articulate the target domain of investigation, and the various properties of the effective domain itself. Consider our finches again: effective laws will only be partially grounded in the universal laws. In addition, they will be partially grounded in the taxon of finches, and in the specifics of their ecological niche.

5. Conclusion

In this article, we've sought to identify a principle which can guide beliefs about physical possibility based on scientific theories. We concluded that current offerings based on laws are too restrictive: they present a rather narrow picture of when the models of theories correspond to physical possibilities. We thus proposed an alternative. On our alternative, it is the models of effective theories that correspond to physical possibilities by approximating parts of worlds.

This alternative is superior for three reasons. First, it does not require theories that capture universal laws. It only requires theories that capture effective laws, which is a less demanding requirement. Second, unlike other options, our principle allows that the models of theories like Newtonian gravitation and evolution by natural selection correspond to physical possibilities. Third, our principle can, with minimal adjustments, accommodate cases in which the models of a theory fail to correspond to any physical possibility, even in part.

Note that this is a first stab at trying to settle on an appropriate principle to guide belief about physical possibility, and so work remains to be done. The principle needs to be tested against a larger range of theories in both physics and science more generally in order to ensure that it continues to deliver plausible results. We also need to consider the principle in light of the full range of ways in which physical possibilities are used in science, to ensure that it can appropriately scaffold scientific practice. It is plausible that the principle will require further refinement going forward. In this way, what we have said provides a foundation for future work on this topic.

5. Compliance with Ethical Standards

- Disclosure of potential conflicts of interest: No conflicts to report
- Research involving Human Participants and/or Animals: N/A
- Informed consent: N/A

5. References

Berenstain, N. and J. Ladyman (2012). Ontic structural realism and modality. In E. Landry and D. Rickles (Eds.), *Structural Realism: Structure, Object, and Causality*, pp. 149–168. Springer.

Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford: Oxford University Press.

Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Earman, J. (1986). *A Primer on Determinism*. Dordrecht: Springer.

Earman, J. (1995). *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes*. New York: Oxford University Press.

Frigg, R. and J. Nguyen (2017). Models and representation. In L. Magnani and T. Bertolotti (Eds.), *Springer Handbook of Model-Based Science*, pp. 25–48. Dordrecht: Springer.

Gell-Mann, M. (1956). The interpretation of the new particles as displaced charge multiplets. *Il Nuovo Cimento (1955-1965) 4*(2), 848–866.

Hoefer, C. (2020). Scientific Realism without the Quantum. In S. French and J. Saatsi (Eds.), *Scientific Realism and the Quantum*, pp. 19–34. Oxford University Press.

Kripke, S. (1981). *Naming and Necessity*. Wiley-Blackwell.

Ladyman, J. and D. Ross (2007). *Every Thing Must Go: Metaphysics Naturalised*. Oxford: Oxford University Press.

Lange, M. (2012). There sweep great general principles which all the laws seem to follow. *Oxford Studies in Metaphysics 7*, 154.

Lange, M. (2016). *Because Without Cause*. Oxford: Oxford University Press.

Lehmkuhl, D. (2017). Literal versus careful interpretations of scientific theories: The vacuum approach to the problem of motion in general relativity. *Philosophy of Science 84*(5), 1202–1214.

Lutz, S. (2017). What was the syntax-semantics debate in the philosophy of science about? *Philosophy and Phenomenological Research 95*(2), 319–352.

Manchak, J. B. (2016). On Gödel and the ideality of time. *Philosophy of Science 83*(5), 1050–1058.

Maudlin, T. (2007). *The Metaphysics within Physics*. Oxford: Oxford University Press.

Pooley, O. (2001). Relationism rehabilitated? II: Relativity. Unpublished manuscript. `http://philsci-archive.pitt.edu/221/`.

Read, J. and N. J. Teh (2022). *The physics and philosophy of Noether's theorems*. Cambridge: Cambridge University Press.

Rivat, S. (2021). Effective theories and infinite idealizations: A challenge for scientific realism. *Synthese 198*(12), 12107–12136.

Ruyant, Q. (2019). Structural realism or modal empiricism? *British Journal for the Philosophy of Science 4*(70), 1051–1072.

Ruyant, Q. (2021). *Modal Empiricism: Interpreting Science Without Scientific Realism*. Springer.

Saunders, S. (2003). Physics and Leibniz's principles. In K. Brading and E. Castellani (Eds.), *Symmetries in Physics: Philosophical Reflections*, pp. 289–307. Cambridge: Cambridge University Press.

Sober, E. (1983). Equilibrium explanation. *Philosophical Studies 43*(2), 201–210.

van Fraassen, B. C. (1980). *The Scientific Image*. Oxford: Clarendon Press.

van Fraassen, B. C. (2002). *The Empirical Stance*. New Haven: Yale University Press.

Williams, P. (2019). Scientific realism made effective. *British Journal for the Philosophy of Science 70*(1), 209–237.

Wilson, A. (2020). *The Nature of Contingency: Quantum Physics as Modal Realism*. Oxford: Oxford University Press.

Wilson, A. (2021). Counterpossible reasoning in physics. *Philosophy of Science 88*(5), 1113–1124.

Woodward, J. (2003). *Making Things Happen*. New York: Oxford University Press.